

Historic, archived document

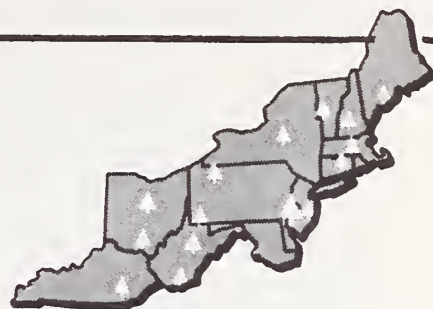
Do not assume content reflects current scientific knowledge, policies, or practices.

1980

A99.9
F76224

Cop. 3

Northeastern Forest Experiment Station



FOREST SERVICE, U.S. DEPARTMENT OF AGRICULTURE, 370 REED ROAD, BROOMALL, PA 19008

EXTRA COPY

TIME SERIES ANALYSIS OF MONTHLY PULPWOOD USE IN THE NORTHEAST

—JAMES T. BONES

Research Forester
Northeastern Forest Experiment Station
Broomall, Pa.

Abstract. Time series analysis was used to develop a model that depicts pulpwood use in the Northeast. The model is useful in forecasting future pulpwood requirements (short term) or monitoring pulpwood-use activity in relation to past use patterns. The model predicted a downturn in use during 1980.

In many decisionmaking situations, past values may be the best variables to use when forecasting future values. Developing stochastic models can be both a necessary and difficult endeavor. Such models are necessary in an industrial society because we live in a world of increasing population and limited natural resources and raw materials. For example, to satisfy regional requirements for paper and allied products, adequate supplies of pulpwood and woodpulp must be forthcoming, or substitute materials must be found. With forecasts of expected pulpwood demand, industrial decisionmakers can devise strategies for obtaining the pulptimber resources that will be needed.

The difficult aspect of developing a predictive model is that variables that are currently known to be highly related to pulpwood use are assumed to be important

prognosticators of future use. In addition to the forecaster's inability to incorporate technological and other change into his predictive model, he is also vulnerable to the mathematical constraints of his model. Simple time series models may only be able to predict broad trends within the current data base, while other types of models with more parameters may account for every aberration. When these models are used to project as many as ten periods into the future, the simple model often proves to be the most helpful because the decisionmaker understands its limitations and is able to allow for possible errors. Although the time series model may predict actual future values closely, it can only forecast future points if the past trends and cycles are repeated. This fact points to another important use of time series analysis; the technique allows the analyst to determine whether

the system he is studying is operating normally. By using his model as a standard for comparison, he can identify abnormal periods.

DATA BASE

For many years the American Pulpwood Association has collected monthly statistics on pulpwood use and ending inventory from all member company pulpmills in the United States and has published monthly summaries by geographic regions (American Pulpwood Assoc. 1977). Pulpwood-use statistics for nine states of the Northeast region¹ between January 1967 and August 1979 were used for developing a time series model.

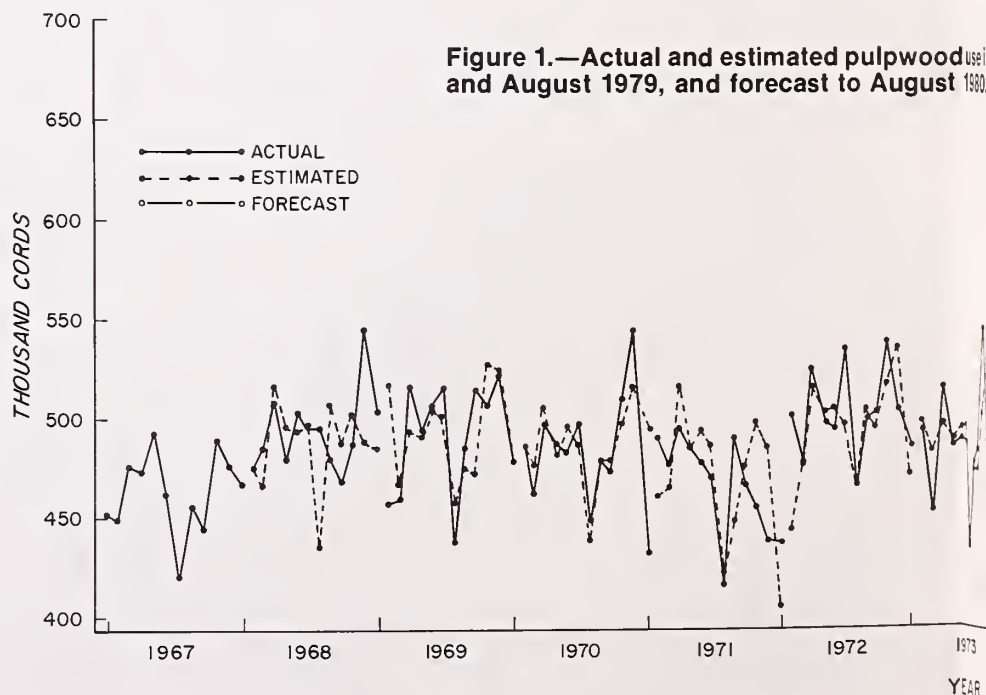
A plot of the regional pulpwood-use statistics revealed that there is a predictable seasonal pattern (Fig. 1). The normal sequence of use in the Northeast has been to start the year with low-use levels. Pulpwood use proceeds to build until peaks are reached in the spring. These peaks are followed by a drop in use as pulpmill workers go on their summer vacations and major mill

maintenance is done. If the demand for woodpulp remains high, low pulpwood use in July is followed by a resurgence in use that extends into the fall. As winter approaches and inclement weather hampers pulpmill operations, pulpwood use once more declines.

METHODS

A number of quantitative forecasting techniques are available for making predictions based on historical analysis; regression (Neter and Wasserman 1974), the U.S. Census Seasonal Adjustment Program (X-11) (U.S. Bureau of Census 1967), and exponential smoothing (Brown 1959) are only a few. Recently Box and Jenkins (1970) proposed a structured approach to time series model building by unifying material and techniques that have been available for a long time. In general, their analysis technique accounts for four components of a series, namely, long-term trends, cyclical effects, seasonal effects, and random variation. The first three components are referred to as the "signal" and the random variation is called "noise." A model is developed which, to the greatest extent possible, eliminates the signal so that only the noise remains. The relative

¹Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont.



contribution of the noise is determined by the autocorrelation function. Based on standard tests such as Bartlett's (1946) approximation or the Q statistic (Box and Pierce 1970), if the model fits the observed data except for random variation, then it is deemed acceptable. A time series model, however, can only be validated when forecasts are compared with actual future values.

The autoregressive integrated moving average process (ARIMA) developed by Box and Jenkins is a powerful tool for modeling time series. Contrary to other model-building techniques, the analyst does not pick a specific model, but instead eliminates inappropriate models until he is left with the most suitable one. MINITAB II (Ryan 1978), a general-purpose statistical computing system, has recently added a computer algorithm for time series analysis based on ARIMA. The user can generate a series of models by specifying the number and kind of parameters. The program provides sample estimates of the model parameters with their standard errors, plots of autocorrelation functions and partial autocorrelation functions, and forecasts of future

observations and their 95 percent confidence intervals.

The preferred model is the one which best represents the data with the fewest number of parameters that are consistent with the stochastic structure. Parameters should be statistically significant and the residuals should not be serially correlated.

RESULTS

The model that best depicts monthly pulpwood use in the Northeast is:

$$(1-B^{12})(1-\phi_1 B-\phi_2 B^2)Z_t = (1-\theta_{12} B^{12})a_t$$

where:

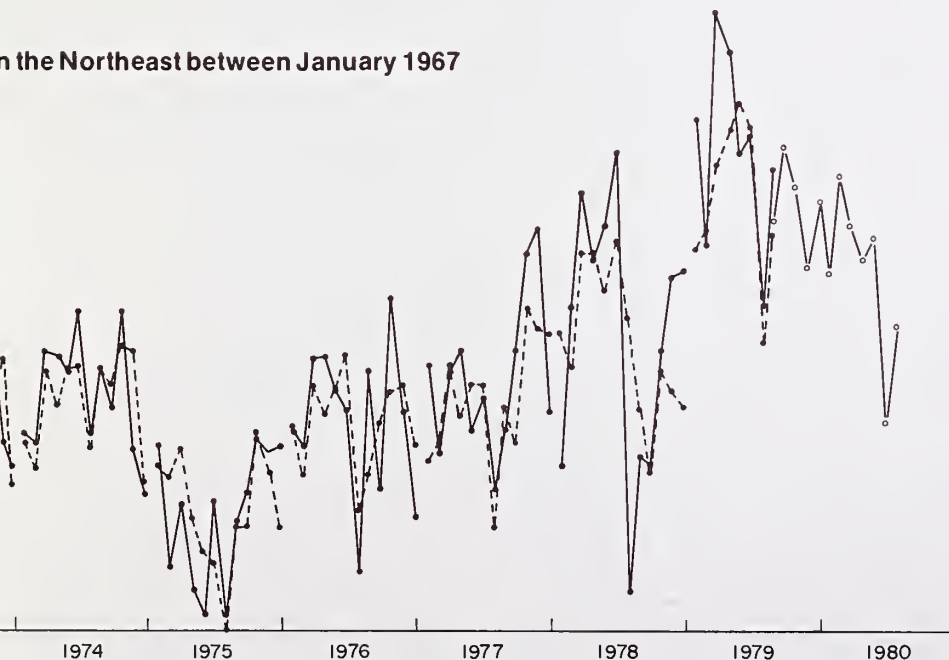
B^m = backward shift operator ($B^m Z_t = Z_{t-m}$)

Z_t = observation at time (t)

a_t = random error at time (t)

The model² has two autoregressive parameters (ϕ_1 and ϕ_2) and one seasonal moving-

²For detailed explanation, see Box and Jenkins 1970, p. 8.





the system he is studying is operating normally. By using his model as a standard for comparison, he can identify abnormal periods.

DATA BASE

For many years the American Pulpwood Association has collected monthly statistics on pulpwood use and ending inventory from all member company pulpmills in the United States and has published monthly summaries by geographic regions (American Pulpwood Assoc. 1977). Pulpwood-use statistics for nine states of the Northeast region¹ between January 1967 and August 1979 were used for developing a time series model.

A plot of the regional pulpwood-use statistics revealed that there is a predictable seasonal pattern (Fig. 1). The normal sequence of use in the Northeast has been to start the year with low-use levels. Pulpwood use proceeds to build until peaks are reached in the spring. These peaks are followed by a drop in use as pulpmill workers go on their summer vacations and major mill

maintenance is done. If the demand for woodpulp remains high, low pulpwood use in July is followed by a resurgence in use that extends into the fall. As winter approaches and inclement weather hampers pulpmill operations, pulpwood use once more declines.

METHODS

A number of quantitative forecasting techniques are available for making predictions based on historical analysis; regression (Neter and Wasserman 1974), the U.S. Census Seasonal Adjustment Program (X-11) (U.S. Bureau of Census 1967), and exponential smoothing (Brown 1959) are only a few. Recently Box and Jenkins (1970) proposed a structured approach to time series model building by unifying material and techniques that have been available for a long time. In general, their analysis technique accounts for four components of a series, namely, long-term trends, cyclical effects, seasonal effects, and random variation. The first three components are referred to as the "signal" and the random variation is called "noise." A model is developed which, to the greatest extent possible, eliminates the signal so that only the noise remains. The relative

contribution of the noise is determined by the autocorrelation function. Based on standard tests such as Bartlett's (1946) approximation or the Q statistic (Box and Pierce 1970), if the model fits the observed data except for random variation, then it is deemed acceptable. A time series model, however, can only be validated when forecasts are compared with actual future values.

The autoregressive integrated moving average process (ARIMA) developed by Box and Jenkins is a powerful tool for modeling time series. Contrary to other model-building techniques, the analyst does not pick a specific model, but instead eliminates inappropriate models until he is left with the most suitable one. MINITAB II (Ryan 1978), a general-purpose statistical computing system, has recently added a computer algorithm for time series analysis based on ARIMA. The user can generate a series of models by specifying the number and kind of parameters. The program provides sample estimates of the model parameters with their standard errors, plots of autocorrelation functions and partial autocorrelation functions, and forecasts of future

observations and their 95 percent confidence intervals.

The preferred model is the one which best represents the data with the fewest number of parameters that are consistent with the stochastic structure. Parameters should be statistically significant and the residuals should not be serially correlated.

RESULTS

The model that best depicts monthly pulpwood use in the Northeast is:

$$(1-B^{12})(1-\phi_1B-\phi_2B^2)Z_t = (1-\theta_{12}B^{12})a_t$$

where:

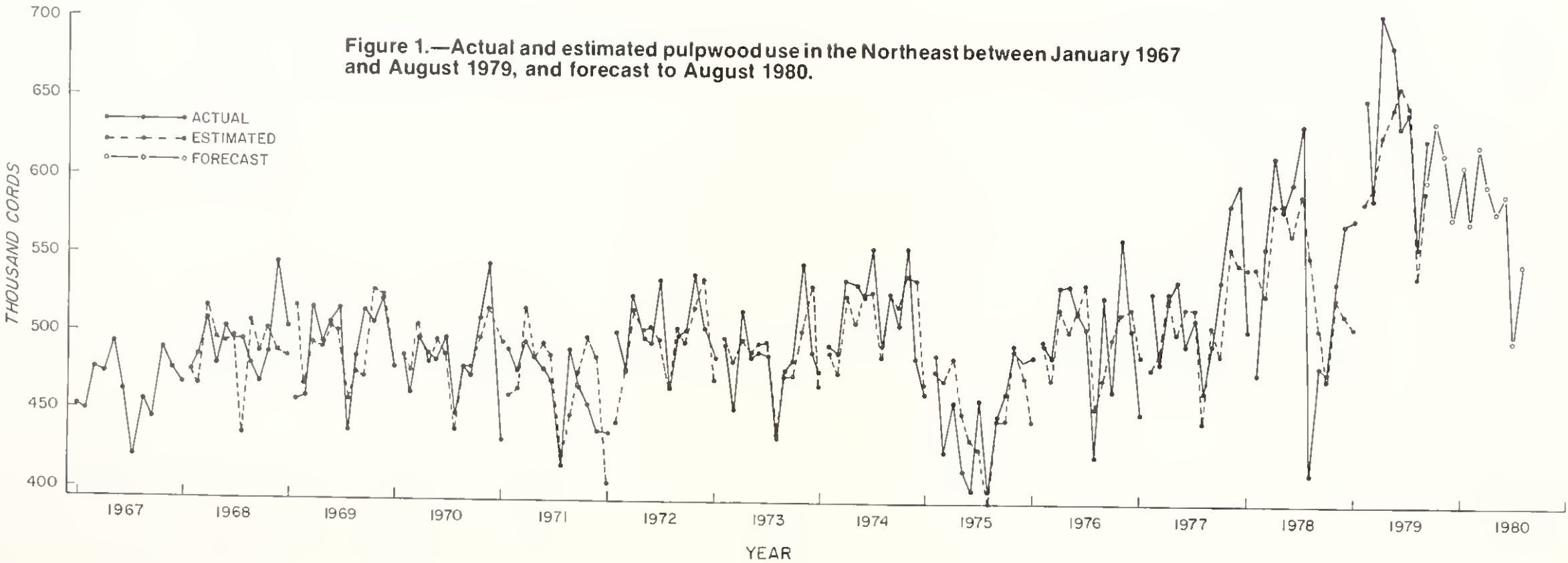
B^m = backward shift operator ($B^m Z_t = Z_{t-m}$)

Z_t = observation at time (t)

a_t = random error at time (t)

The model² has two autoregressive parameters (ϕ_1 and ϕ_2) and one seasonal moving-

²For detailed explanation, see Box and Jenkins 1970, p. 8.



¹Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont.

average term (θ_{12}). Differencing was required for the seasonal portion of the model. The estimate of parameters and their standard errors are:

<i>Parameter</i>	<i>Estimated value</i>	<i>Standard error</i>
ϕ_1	0.6085	0.0835
ϕ_2	0.2617	0.0853
θ_{12}	0.8210	0.0750

The pulpwood-use model was judged satisfactory in that all of the parameter estimates were statistically significant, and none of the sample autocorrelations were statistically significant so that the residuals were not autocorrelated.

Based on this model, Figure 1 shows a comparison of the actual pulpwood-use statistics and the predicted values for the Northeast between January 1968 and August 1979. In addition, pulpwood use is forecast to August 1980. During the 11-year period, all of the predicted values were within ± 13 percent of the actual values, except in July 1978 when actual pulpwood use dropped 25 percent below predicted use. This illustrates how the model can be used as a means of identifying abnormal levels of use.

DISCUSSION

By using the ARIMA technique, a time series can be fitted with a mathematical model that is optimal in the sense that it assigns smaller errors to history than any other model. The type of model is identified and the parameters are then estimated. This statistical routine is currently considered to be one of the most accurate portrayals of time series for short-term forecasting. Although forecasts are used by resource and raw material planners to estimate future requirements, they are also used by economists to gauge market supplies and demands relative to the norm. Calendar year 1975 was an atypical year for pulpwood use in the Northeast. Forecasts based on time series analysis data between 1967 and

1974 tracked monthly use of between 450 and 550 thousand cords. Actual monthly use, however, dropped to between 400 and 450 thousand cords during 1975. As a result, there were pulpwood inventory build-ups in many northeastern woodyards. Additional wood inventories tie up working capital and increase the risk of wood deterioration during the longer holding period.

Our time series model forecasts that 7.0 million cords of pulpwood will be required in the Northeast between August 1979 and August 1980. This figure represents a 1.0 percent decrease in use from the same period in 1978-1979.

LITERATURE CITED

- American Pulpwood Association.
1977. **Pulpwood statistics.** Am. Pulpwood Assoc., Washington, D.C.
- Bartlett, M.S.
1946. **On the theoretical specification of the sampling properties of autocorrelated time series.** J. R. Stat. Soc. B8:27.
- Box, G. E. P., and G. M. Jenkins.
1970. **Time series analysis, forecasting, and control.** Holden-Day, Inc., San Francisco.
- Box, G. E. P., and D. A. Pierce.
1970. **Distribution of residual autocorrelations in autoregressive moving average time series models.** J. Am. Stat. Assoc. 65(332): 1509-1526.
- Brown, R. G.
1959. **Less risk in inventory estimates.** Harv. Bus. Rev. July-Aug. p. 104.
- Neter, J., and W. Wasserman.
1974. **Applied linear statistical models.** Richard D. Irwin, Inc., Homewood, Ill.
- Ryan, T. A., Jr.
1978. **Minitab II.** Pa. State Univ. Comput. Cent., University Park. Release 78.1.
- U.S. Bureau of Census.
1967. **The X-11 variant of the census method I seasonal adjustment program.** Tech. Pap. 15 (rev.) U.S. Dep. Commer., Bur. Census. Washington, D.C.